DOCUMENT RESUME

ED 049 310 TM 000 510

AUTHOR Blumenfeld, Gerald J.; And Others
TITLE Effect of Criterion Referenced Testing Upon the Use of Remedial Exam Opportunities.
PUB DATE Feb 71
NOTE 10p.; Paper presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS College Students, *Comparative Testing, *Criterion Referenced Tests, *Experimental Teaching, Instructional Programs, *Remedial Instruction, Scoring, *Student Testing, Testing Programs

ABSTRACT

Groups of approximately 300 undergraduate educational psychology students participated in an instructional program, involving weekly instructional objectives, weekly exams, remedial exams, and a point system of scoring. It was assumed that the percentage of students not attaining the passing criterion on the initial exam but returning for the remedial exam that week would be greater for a high criterion condition, and that the high criterion condition would also result in a greater percentage of students attaining the criterion. A follow-up study, with two groups assigned to alternating conditions, was conducted. Class and absolute point systems of scoring were used to determine letter grades. Further replication of the study substantially confirmed the initial findings. The hypothesis was supported by the results and it was concluded that the provision of optional remedial opportunities would have greater effects on student utilization and achievement if relatively high criterion-referenced testing programs were used concurrently. (LR)

ED049310

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Effect of Criterion Referenced Testing Upon the Use of Remedial Exam Opportunities

Gerald J. Blumenfeld, University of Akron
Darrel Bostow, University of South Florida
Robert Waugh, Southern Illinois University

Consideration of norm and criterion referenced testing (Glaser, 1963) has become a potentially significant topic in the area of measurement and evaluation. While discussion concerning the technical aspects of constructing criterion referenced tests continues (Popham, 1969) various articles have reported instructional programs which incorporate the idea of criterion referenced measurement (Keller, 1968; Ferster, 1968). As As with many other instructional procedures, the effect of this variable upon student behavior and the interaction it may have with other instructional variables is yet to be determined.

During the winter quarter of 1969, the undergraduate educational psychology class at SIU Carbondale was reorganized by the senior author and began registering students for sections of approximately 300. All sections employed an instructional program characterized by weekly instructional objectives, weekly exams, remedial exams, and a point system. Specific details varied according to the disposition of the instructors teaching a section. One section imposed a rather high criterion (85%) for receiving passing credit for weekly exams. In this section 95% or more of the students not meeting the criterion on the initial test would return for the remedial exam. Approximately 90% of the students in this section received a grade of "A" for the course. In sections that did not impose such a high criterion only 25% of the students who could profit from the remedial exam would return. About 30% of the students in these

TM 000 510

AERA, N.Y., 1971 C-76

sections received a grade of "A" for the course. Since these data were gleaned from records involving different instructors, objectives, exams, students, and quarters, it was decided to collect such data in a more rigorous manner. It was assumed that the percentage of students not attaining the passing criterion on the initial exam but returning for the remedial exam that week would be greater for a high criterion condition, and that the high criterion condition would also result in a greater percentage of students attaining the criterion.

Method

During the winter quarter of 1970, one section of the course was divided into two groups. The groups were assigned to alternating conditions which varied the criterion required for receiving passing credit on the weekly exams composed of 20 multiple choice items. Each student selected his seat for the term during the second meeting of the class. After seats had been selected students were informed that rows A, C, E, etc., were in Group I, and rows B, D, F, etc., were in Group II.

All students received a semi-programmed explanation of the course organization during the first week and were not given credit for other work until they had answered correctly 18 of 20 items on a test covering the organization. Only seven students failed to pass this exam the first time. The program informed students of the weekly routine and of the two point systems. Monday would involve lectures or films; Tuesday, the initial exam and distribution of objectives for the following week; Wednesday, no class; Thursday, a remedial lecture and discussion period; Friday, the remedial exam. Attendance was never required.

All exams were projected using overhead transparencies prepared with large primary type. Two sequences of the same items were projected

simultaneously. Students assigned to odd numbered seats answered the items projected onto the left screen; students assigned to the even numbered seats answered items appearing on the right screen. Each student kept a record of his answers on a one inch strip of colored paper he retained. In addition he recorded answers on a docutran sheet which was scored by an optical scanner. Immediately after collecting the docutran sheets the instructors either projected or distributed a scoring key. An attempt was made to facilitate study for the remedial exam by identifying the objective that each item was intended to assess. Test items were not made available for study, and remedial exams contained at least 50% new items. Each student's official score was printed on a computer listing and posted before class on Thursday.

One scoring procedure was called the class point system. Under this condition initial and remedial exams were scored as follows:

Initial Test		Remedial Test	
raw score	class points	raw score	class points
16 - 20	4	17 - 20	3
11 - 15	1	14 - 16	1
10 or less	0	13 or less	0

Points from the initial and remedial exams were additive, but four was the maximum number of class points awarded for any one week.

The second scoring procedure, called the absolute point system, credited each student with his raw score. If he came back for the remedial exam his score for the week was the higher score of the two exams. At the end of the term all class points accumulated were multiplied by five and added to the points accumulated from raw scores under the absolute point system. A letter grade was assigned using the following scale:

Total Points Accumulated (class points X 5 plus raw score points)	Letter Grade
155 to 180	A
135 to 154	B
115 to 134	C
95 to 114	D
below 95	F

The final grade was raised one letter if the student's score on a required comprehensive final was 80% or more and lowered one letter if the final exam score was less than 60%.

Beginning with the first unit exam administered during the second week, the two groups were assigned to the two criterion conditions as follows:

Week	Unit	Group I	Group II
2	1	class points	absolute points
3	2	class points	absolute points
4	3	absolute points	class points
5	4	absolute points	class points
6	5	class points	absolute points
7	6	class points	absolute points
8	7	absolute points	class points
9	8	absolute points	class points

No student was allowed to take the remedial exam unless he had taken the initial exam. Students were excused for missing the initial exam and permitted to make it up only if they had a legitimate reason that was documented. (Unit 9 not included in study.)

Results

Figure 1. represents the percentage of students not attaining the 80% criterion on the initial exam for each unit but returning to take the remedial exam. On the average 50% of the students attained a score of 16 or more on the initial exam. Figure 1. indicates that the differences were consistent and that group or order effects did not exist. Using a sign test the probability of obtaining 8 of 8 comparisons

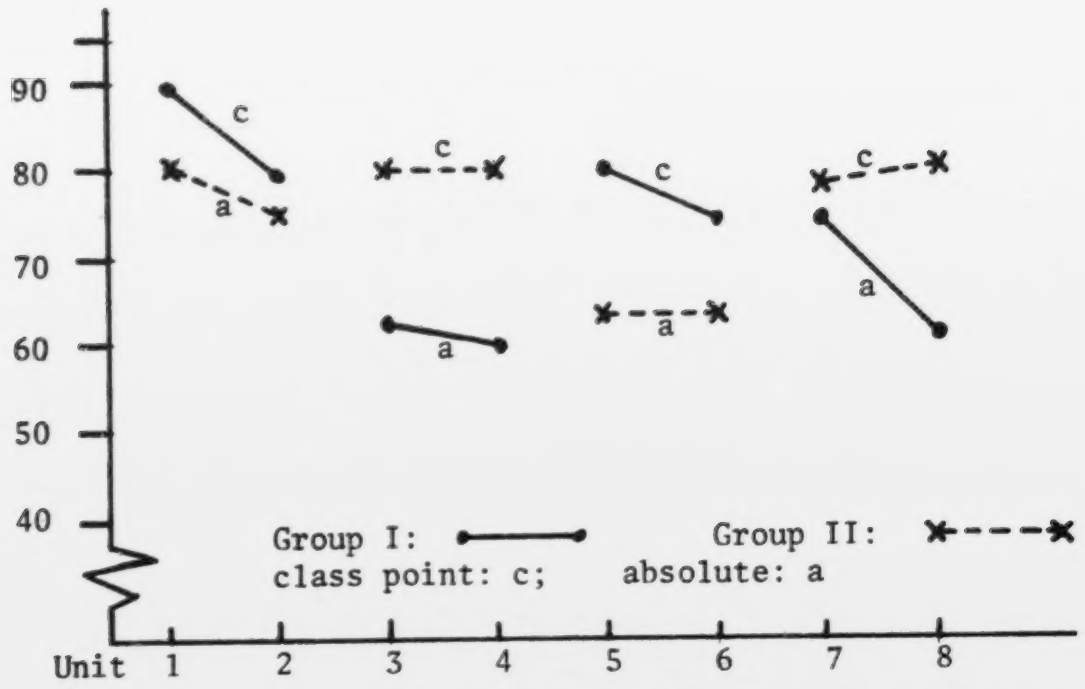

Figure 1. Winter quarter. Percent of students scoring below criterion on initial exam who took remedial exam.

in the predicted direction is .004.

Figure 2. indicates the percentage of students attaining either 80% on the initial exam or 85% on the remedial exam. Using a sign test the probability of obtaining 7 of 8 comparisons in the predicted direction is less than .035. The one difference not in the predicted direction occurred during the first unit before students had any personal experience with the criterion policy.

Replications

Attempts to replicate the study were made during the spring and summer sessions of 1970. In both cases the same instructors and procedures were used. Data for the first five weeks of the spring quarter was obtained and is reported in Figures 3 and 4. It is in substantial agreement with that collected during the winter even though

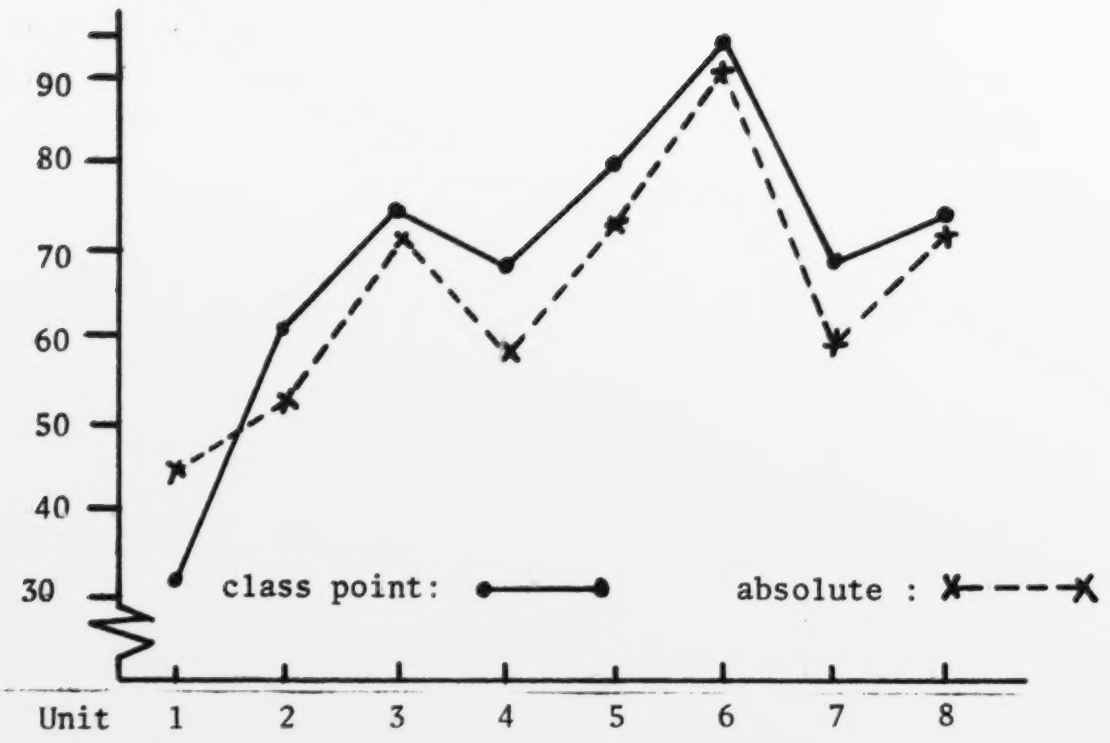

Figure 2. Winter quarter. Percent of students who attained criterion on either initial or remedial exam.

the campus and civil disorders that eventually resulted in closing the school on May 12th disturbed normal routines for nearly half the time school was open that quarter.

Data for the summer session is reported in Figures 5 and 6. One hundred sixty one students were involved in the study during the summer rather than 300 as in other quarters. Approximately 40 additional students enrolled for the summer volunteered for a special readings section of the course and were not included in the study. Volunteers were solicited during other quarters, but the percentage of volunteers taken out of the large section was quite small. The summer data is also consistent with that obtained during the winter. Seven of the eight comparisons involving percentage of students who returned to take the remedial exam were in the predicted direction. Six of the eight comparisons involving percentage

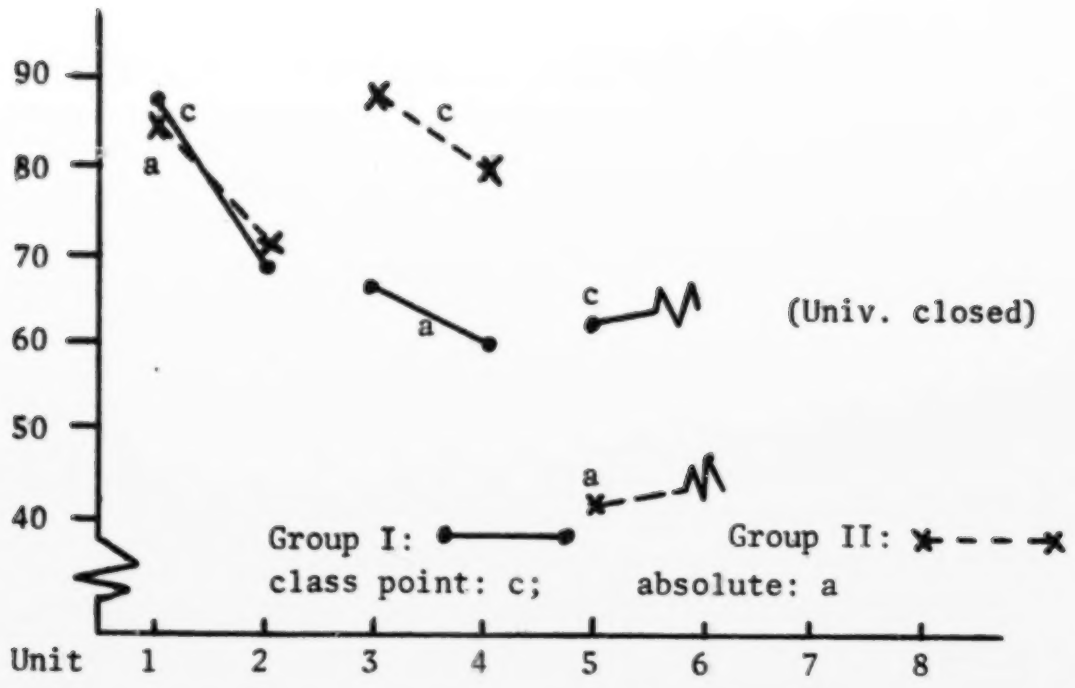

Figure 3. Spring quarter. Percent of students scoring below criterion on initial exam who took remedial exam.

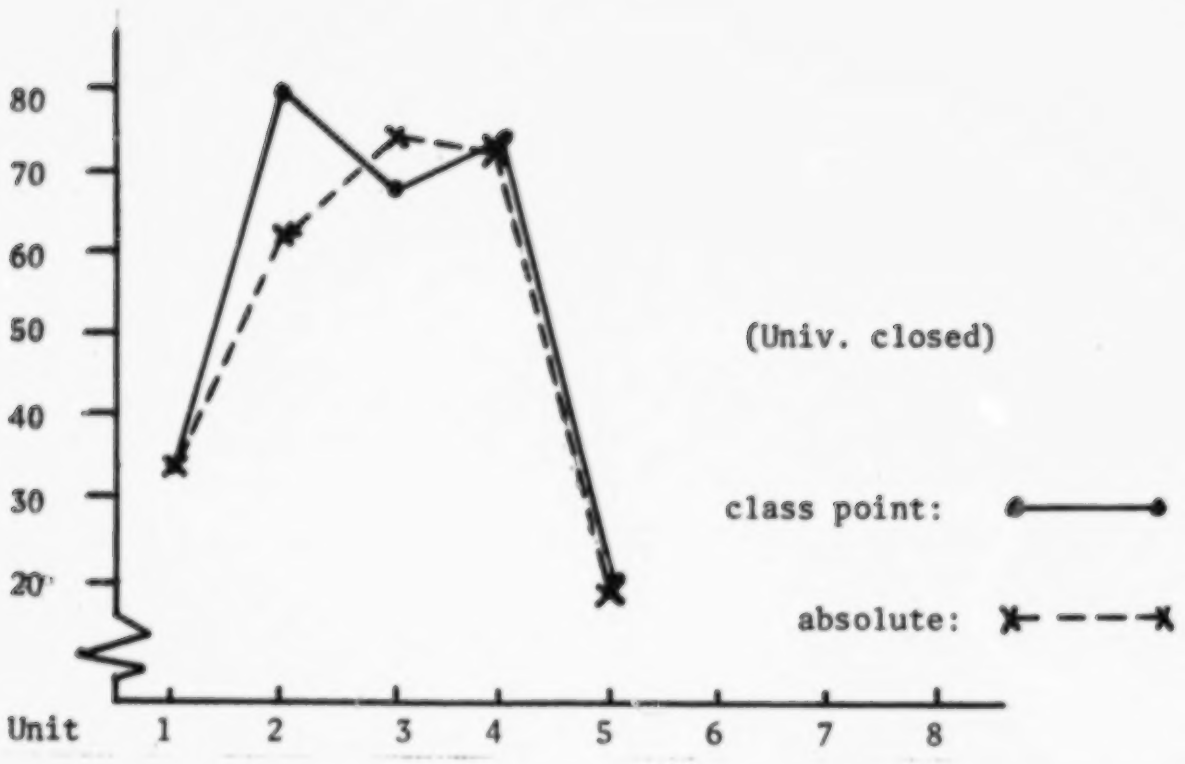

Figure 4. Spring quarter. Percent of students who attained criterion on either initial or remedial exam.

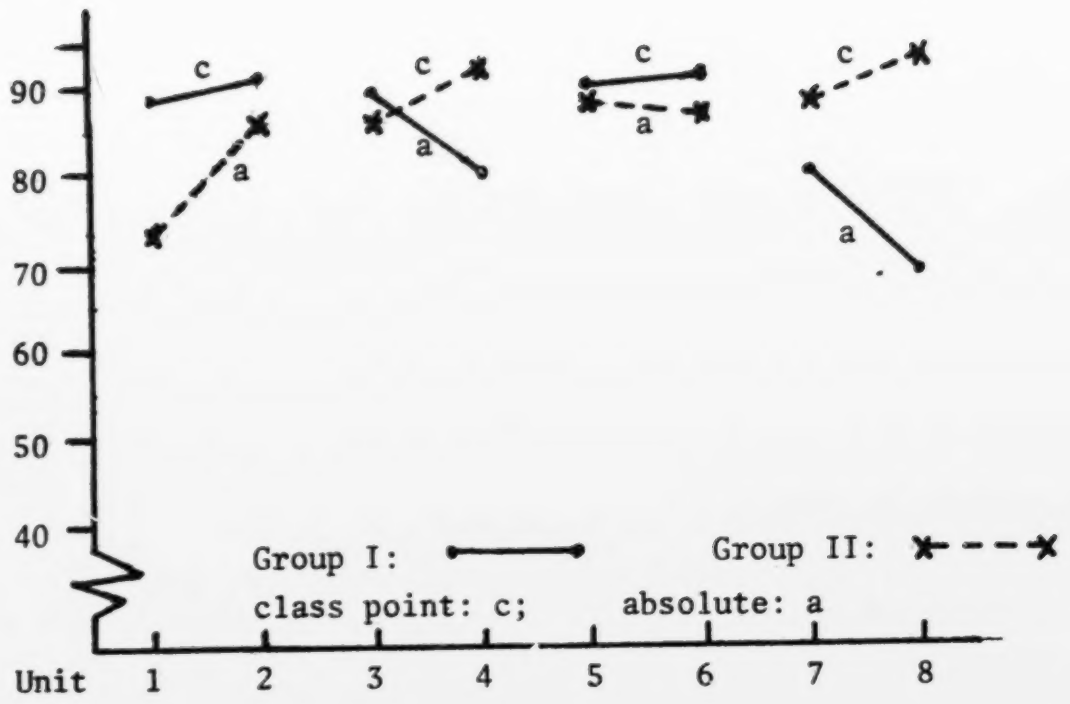

Figure 5. Summer quarter. Percent of students scoring below criterion on initial exam who took remedial exam.

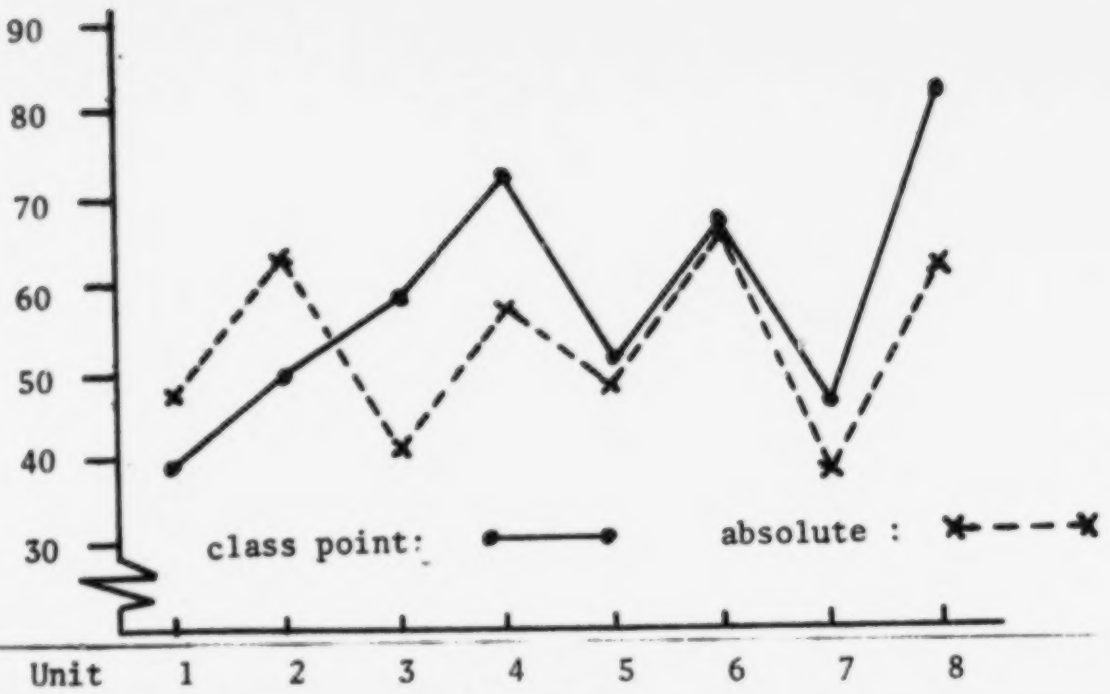

Figure 6. Summer quarter. Percent of students who attained criterion on either initial or remedial exam.

of students attaining the criterion on either the initial or remedial exam were in the predicted direction. It is interesting to note that once again the reversals that did occur were in the first two weeks.

Discussion

The consistency in the direction of the differences suggests rather strongly that the differences are reliable ones. However, it is also interesting that each time the data was collected the percentage of students using the remedial opportunity increased for those not required to attain a high criterion. One could assume that exposing a student to both conditions would moderate the effects. The winter data is consistent with this assumption. However, if such a moderating effect exists it is not evident in the summer data. The percentage of students using the remedial exams during the summer was close to that obtained when a high criterion was imposed for the entire term. This could have been the result of a different population. It is possible that the summer students were more studious. Or one might assume that the disruptions on campus during the spring tended to suppress non-studying behaviors during the summer. It is also possible that operating with remedial exams for a year and a half resulted in students learning about the exams and beginning the summer course with the expectation of taking advantage of such opportunities.

Altogether, 19 out of 21 comparisons indicate that a larger percentage of students who could profit from a remedial exam take that exam when a high criterion is imposed. Sixteen out of 21 comparisons indicate that a larger percentage of students attain a high criterion when only minimal credit is given for lower scores. These data suggest that the effort and expense required to provide optional remedial oppor-

tunities may have a greater effect upon student utilization and achievement if relatively high criterion referenced testing programs are employed concurrently.

Bibliography

Ferster, C.B. Individualized instruction in a large introductory psychology college course. Psychol. Rec., 1968, 18, 521-532.

Glaser, R. Instructional technology and the measurement of learning outcomes: Some questions. Amer. Psychol., 1963, 18, 519-521.

Keller, F. "Good-Bye, Teacher . . ." J. applied Behav. Anal., 1968, 1, 79-89.

Popham, W.H., & Husek, T.R. Implications of criterion-referenced measurement. J. Educ. Meas., 1969, 6, 1-9.

Simon, G.B. Comments on "Implications of criterion-referenced measurement." J. Educ. Meas., 1969, 6, 259-260.